MERCHANDISING MEASURES FOR A PRODUCT LINE

John D.C. Little

School of Management
M.I.T.
Cambridge, MA 02139 U.S.A.

May 1989 WP# 3024-89-MS

ABSTRACT

Managers look at product line data on sales, distribution, and merchandising and draw conclusions about marketing effectiveness. Therefore measures of these variables for a product line should be as meaningful as possible. We propose that a good measure is one whose changes drive changes in sales. By theoretical analysis and empirical testing of a simple merchandising model, we show that useful measures of merchandising and distribution for a product line can be constructed by taking weighted averages of the data for individual items. The weights are base sales rates at full distribution. An example from Infoscan data for the Ocean Spray Cranberry Juice Cocktail line shows that the proposed measures are better than those currently used.

1. Introduction

Much of the excitement over single source data for consumer package goods springs from its remarkable detail. Bar code scanners at supermarket checkout individually record each package and its price. Data companies supplement this with other observations in the store and in the market so that, today, we can have at our fingertips sales, distribution, display, advertising and other merchandising activity by individual item, week, retailer, and market.

Despite this detail, however, managers still need summary measures that express, in a sensible way, what is happening to a brand or a product line as a whole over broad geographic regions and time periods. We observe, in fact, that managers look at the aggregate measures, draw conclusions from them, and make decisions based on back-of-the-envelop calculations with them, whether marketing analysts like it or not. It behooves analysts and data companies, therefore, to provide the most meaningful aggregate measures they can. Although our focus will be on consumer package goods, where large databases have forced the issue, the ideas to be discussed apply more generally.

Three kinds of aggregation are usually sought: totals over geographic areas, time periods, and products. Commonly, data is collected from a sample of retail stores and used to represent a geographic area. Projection to the area is done by weighting stores according to their size as measured by all commodity volumes (ACV). This procedure is appropriate and will be assumed. Aggregation over time can usually be done by adding (or averaging) over time periods. We shall argue that this should be done in certain cases where it is not done now. Summarizing measures over products is a less obvious process and will be the main subject of our paper.

Sales are usually easy to aggregate, even when items have different sizes. Most manufacturers define a unit, such as "equivalent volume" to make different sizes comparable. Then equivalent volumes can be added to determine meaningful totals for a product line. Adding sales across time periods is also straightforward.

Much more troublesome, however, are the current conventions for distribution and merchandising. The two leading data sources for the consumer package goods industry, Information Resources Inc.'s (IRI) Infoscan and A. C. Nielsen's Scantrack, both use "non-additive" measures for distribution and merchandising. See, for example, IRI (1988). Table 1 shows a week's data for an aggregate product line, Ocean Spray Cranberry Juice Cocktail, and the eight individual Ocean Spray items that make it up. Shown are the distribution and display measures for the individual items and, as currently reported, for the line as a whole.

Distribution for an individual item is the % of stores with non-zero sales of the item, where each store is weighted by its size expressed as its all commodity volume, and so has units of % ACV. For example, the 74 points of ACV distribution for Cranberry Juice Cocktail Liquid Concentrate in Table 1 mean that this product was sold during the week in a set of stores that constitute 74% of the all commodity volume in the geographic area, which in this case is the total US. Display (or other merchandising measure) has a similar meaning: The value of 2.1 % ACV on display for Cranberry Juice Cocktail Aseptic 3-Pack means that a set of stores representing 2.1 % of the all commodity volume in the total US had a display of the item during the week.

The currently used rule for determining the display of a product line like Total Cranberry Juice Cocktail is as follows: The product line is considered to be on display in a store for the week if any of its component items is on display. Similarly, the product line is in distribution in a store if any of its components is in distribution. Thus, in Table 1, we see that Total Cranberry Juice Cocktail has 100% distribution in the week even though none of its individual items is in all the stores. Similarly, the display measure for the aggregate product is much larger than any of its individual items.

Product	Distribution (% ACV)	Display (% ACV)
Cranberry Juice Cocktail Liquid Concentrate	74	.4
Cranberry Juice Cocktail Aseptic 3-pack	69	2.1
Cranberry Juice Cocktail 32 oz. Bottle	97	3.2
Cranberry Juice Cocktail 48 oz. Bottle	99	5.8
Cranberry Juice Cocktail 64 oz. Bottle	95	.3
Cranberry Juice Cocktail 128 oz. Bottle	88	.3
Cranberry Juice Cocktail Low Calorie 32 oz.	48	.4
Cranberry Juice Cocktail Low Calorie 48 oz.	79	2.9
Total Cranberry Juice Cocktail	100	12.9

Table 1. Under currently used rules, an aggregate product like Total Cranberry Juice Cocktail is said to be in distribution (or on display) in a store if any of its component products is in distribution (on display). (Data is for Ocean Spray products for the total US in the week starting February 22, 1987 as reported by IRI's Infoscan.)

A similar rule is used for time aggregation. Consider, for example, a 12-week period. The 32 ounce bottle will be said to be on display in a store during the period if it was on display during at least one of the 12 weeks.

While these measures tell something about what is going on, they quickly saturate to 100% for large aggregates and, a more serious problem, are poor indicators of the depth and strength of distribution and merchandising activity within the product line and over time.

1.1 What properties should aggregate measures have?

We suggest that aggregate measures of merchandising and distribution should, if posssible:

> Summarize activity and permit comparisons across multiple markets (or other collections of stores such as key accounts), time periods, and products (both individual items and product lines).
>
> Represent fundamental variables that drive sales.
>
> Mean the same thing for a product line as for an individual item.

Measures satisfying these criteria will be useful for monitoring and evaluating sales and marketing programs.

1.2 Approach to constructing new measures

As discussed above, we assume that managers (and marketing analysts), look at the aggregate numbers and draw conclusions from them. One might say that they run mental regressions. Therefore, measures of causal variables will be most meaningful if changes in them imply changes in performance variables like sales and share. A good measure is one that would be useful for building market response models, i.e., might reasonably be used to run real regressions.

The analytic approach taken is as follows:

(1) We hypothesize a merchandising response model for an individual product at the store level. The model relates sales to merchandising variables and includes a fundamental response constant for the merchandising activities. The constant is one that could be estimated from scanner data.

An individual product will be defined by its Universal Product Code (UPC), which represents the finest level of detail collected by the scanners.

(2) We apply sampling projection factors to determine what the model says about individual UPC's at the market level.

(3) Next we create product aggregates at the market level and ask: "What are the aggregate merchandisng measures that

would make it possible to estimate the UPC-store level response constant from product aggregate-market level data.

(4) We test the theory using IRI InfoScan data. We estimate the response constant from the individual UPC data, by means of time series regression of sales against merchandising. Then we estimate the response constant again, but now from product aggregate data with the aggregate merchandising measures developed from the theory. The two estimates of the response constant are compared.

We would not expect the estimates to be exactly the same. (In fact it would be an alarming day for the data business if the response constant did not change somewhat with product and market!) However, the estimates turn out to fall within a range that would be expected, thereby indicating that the proposed measures are doing what we want. That is, changes in these measures drive changes in the sales of the product aggregate.

(5) Finally, we compare the proposed measures with those currently being used with respect to their ability to explain and interpret sales performance. The new measures are better.

Here is a brief summary of the new measures. We assume that base sales is available at the market level for each UPC. Base sales is the sales that would take place in the absence of any merchandising activity.

(1) Merchandising (for a given week).

The merchandising measure for a product aggregate is the weighted average of the measures for the individual UPC's, where the weights are the base sales of the UPC's at full distribution. Base sales at full distribution is obtained by dividing the base sales of a UPC by its distribution. This quantity can also be described as the base sales per point of distribution.

The new merchandising measure will have a value of 100 % if and only if all products in the aggregate have the merchandising in all stores of the market during the given week. The value will be correspondingly less for the ususal case in which some of the products do not have merchandising.

(2) Distribution (for a given week).

Similarly, the distribution of a product aggregate is the weighted average of the ACV distributions of its individual UPC's, where the weights are the individual base sales at full distribution.

Distribution will have a value of 100 % if and only if all products in the aggregate have distribution (non-zero sales) in all stores in the market during the week.

(3) Merchandising (over multiple weeks)

This is a simple sum of the measure in each week and represents ACV-weeks of merchandising.

(4) Distribution (over multiple weeks)

Distribution for a multi-week time period is the average over the weeks of interest.

We shall argue that, even if base sales is not available, we can approximate it or replace it with a surrogate with relatively little loss in quality.

2. Theory Development

To simplify the discussion, consider the following to be held fixed:

Time period (a particular week)
Market
All merchandising variables but one.

These restrictions can be relaxed. In particular, we note that, instead of market, we can equally well deal with any population of stores which is being sampled and then projected. For example, we could consider key accounts or sales territories. The merchandising variables to be considered are display, features, and price reductions. Each is either 0 or 1 (absent or present) for any given UPC in a specific store for the week being considered. These variables come in various forms, such as feature only, display only, feature and display, etc. The theory applies to any of them. We can also consider a linear combination of merchandising variables, provided only that it takes on the value zero when the product is not present in the store.

2.1 Store level measures

Here is notation to describe the basic data collected at the store level. Let

s_{ur} = sales rate of UPC u in retail store r (equivalent units/week)

m_{ur} = 1 if merchandising variable m of UPC u is present in store r during week.

0 otherwise.

R = {r} = set of stores in sample.

a_r = all commodity volume (ACV) of store r (millions of dollars/year).

$a_R = \Sigma_{r \epsilon R} a_r$ = ACV of sample R.

a = ACV of market area (millions of dollars/year).

d_{ur} = 1 if UPC u is in distribution in store r during week,

0 otherwise.

In addition some data suppliers calculate

s_{0ur} = base sales of u in store r, i.e., sales rate in the absence of merchandising (equivalent units/week).

This quantity is not directly observable but is obtained by applying one of a number of possible models and algorithms to determine a baseline of sales that would have occurred without promotional activity (Abraham and Lodish, 1987). Such algorithms have become highly developed and are now applied at the individual store level.

Data companies do not report store level data because of confidentiality requirements in their contracts with retailers. Reported are projections to market level. We now express these in terms of the store data.

2.2 Market level measures

Relevant measures collected by the data companies and supplied to manufacturers include the following:

s_u = sales rate of u in the market (equivalent units/week),

m_u = fraction of market ACV with merchandising variable m present for u,

d_u = fraction of market ACV having u in distribution,

and often

s_{0u} = base sales rate of u (equivalent units/week).

These measures are calculated by projecting individual store data to market level as follows:

$$s_u = (a/a_R)\ \Sigma_{r\epsilon R}\ s_{ur}$$

$$m_u = \Sigma_{r\epsilon R}\ (a_r/a_R)\ m_{ur}$$

$$d_u = \Sigma_{r\epsilon R}\ (a_r/a_R)\ d_{ur}$$

$$s_{0u} = (a/a_R)\ \Sigma_{r\epsilon R}\ s_{0ur}$$

Merchandising and distribution in these formulas are dimensionless numbers with a range of 0 to 1. Standard practice, however, is to report them on a scale of 0 to 100 percentage points. Thus, $d_u = .73$ is usually described as 73 points of ACV distribution. We frequently use the percent terminology in discussion, but all equations deal with fractions.

Finally, it will turn out that a key quantity for our work will be base sales at full distribution:

$v_{0u} = s_{0u}/d_u$ = base sales of u at full distribution in the market (equivalent units/week).

This important measure tell us the inherent consumer strength of the product independent of its level of distribution and merchandising activity. Algebraically, it can be shown from the definitions of s_{0u} and d_u to be equivalent to taking the base sales rate in the sample stores where u is present and projecting it to the full market:

$$v_{0u} = (\Sigma_{r\epsilon R}\ s_{0ur})\ (a/\Sigma_{r\epsilon R}\ a_r\ d_{ur})$$

This quantity, when divided by 100, is also called base sales per point of distribution. Packaged goods marketers can use v_{0u}, or any similar measure such as base share per point of distribution, to rank the items within a product line. An individual product that ranks highly is doing well with the customers in the stores that carry it. This means that, if it has low distribution, it will be a strong candidate for sales and marketing effort to increase the number of stores stocking it.

We should point out that, by calling v_{0u} "base sales at full distribution," we are implicitly suggesting a direct proportionality between distribution and sales. This is undoubtedly not quite true for several reasons. For one thing, some customers may go out of their way to shop at a store because it carries the product. For another, the first few stores to carry a product as distribution increases are likely to be qualitatively different from later participants. Nevertheless, there is no question that, as a first order effect, more distribution means more sales and that v_{0u} provides a key measure of inherent product strength.

3. A Merchandising Model

To capture the effect of merchandising in as straightforward a manner as possible, we assume that it produces a simple percentage effect in a basic store sales rate. The percentage is taken to be approximately the same for all stores and UPC's in the category and is expressed in terms of an underlying response constant. In words:

Store level: one product
sales = (base sales) (1 + k merchandising)
k = response constant

In symbols:

$$s_{ur} = s_{0ur}\,[1+km_{ur}] \tag{1}$$

where

s_{0ur} = sales of u in store r in the absence of merchandising m (equivalent units/week).

k = response constant for merchandising m (dimensionless).

m_{ur} = 1 if merchandising variable m of UPC u is present in store r during week.

0 otherwise.

Another way of describing (1) is to say that merchandising of a given type produces different amounts of incremental sales for different products in different stores but percentage-wise the increments are similar.

We now ask: How should (1) be calibrated? One way would be run time series regressions on store level data. We know sales, s_{ur}, and merchandising, m_{ur}, and could estimate base sales, s_{0ur}, and the response constant, k. This would be a good thing to do if we wanted to know base sales and response by individual store. And, as mentioned earlier, the data companies, using more sophisticated models, do this now.

Our goal here is different. We are interested not only in what is happening in stores where the product has distribution but also what might happen in stores where it does not. Thus, a statement like: "Our 32 oz. size has distribution in 80% of ACV," is valuable partly because it tells us that the product is not available in 20% of the market and carries the implicit promise that we would obtain more sales if we could obtain more distribution. In fact, the statement implies that, if we could get 100% distribution, sales should be about 25% (=20/80) higher. This is an important type of thinking that we wish to capture with our measures.

To do this we model s_{0ur} for an arbitrary store by using data from all the stores that carry u. Specifically, noting that

$v_{0u} = s_{0u}/d_u$ = average base sales per unit of ACV distribution among stores carrying u (equivalent units/week),

a_r/a = fraction of market ACV represented by store r,

d_{ur} = 1 if store carries u,

0 otherwise,

we take:

$$s_{0ur} = v_{0u}(a_r/a)d_{ur}. \qquad (2)$$

Another way of describing (2) is that we are estimating base sales in an arbitrary store based on its size, whether or not it carries the product, and the average sales across all stores (weighted for size) that carry the product.

Now we ask: if (2) is the calibration at the store level, what response does it predict at the market level and do the data companies report a merchandising measure tht could be used to estimate k?

Using (1), (2) and earlier definitions,

$$s_u = s_{0u} + k\ (a/a_R)\ v_{0u}\ \Sigma_r(a_r/a)m_{ur}d_{ur}$$

Note that $m_{ur}d_{ur} = m_{ur}$, since $m_{ur} = 0$ whenever $d_{ur} = 0$ and is 1 otherwise. Simplifying gives

$$s_u = s_{0u}\ [1 + k\ (m_u/d_u)] \qquad (3)$$

In words:

Market level: one product

sales = (base sales) (1 + k [merchandising / distribution])

We observe that base sales at the market level depends on distribution. Equation (3) can be written in terms of the more fundamental measure of product strength, v_{0u}:

$$s_u \quad = \quad v_{0u}\, d_u\, [1 + k\, (m_u/d_u)] \qquad (3a)$$

Data companies provide the needed variables to calibrate (3): m_u is the fraction of ACV with merchandising, and d_u is the fraction of ACV with distribution. Therefore we can, if we wish, run a time series regression of s_u against m_u/d_u at the market level and estimate the response constant k. We could also estimate s_{0u}, the market sales volume in the absence of the merchandising in this manner. If distribution for product u varies appreciably over the time period under consideration, the alternate (3a) can be used, regressing s_u/d_u against m_u/d_u.

Finally we note that we can interpret the store level model (1) as being the same as (3) at the market level, since, at the store, if the product is present, its distribution is 1.0. Or, more fundamentally, since v_{0ur}, base sales at full distribution, equals s_{0ur} if the product is present, we can rewrite (1) as

$$s_{ur} \quad = \quad v_{0ur}\, d_{ur}\, [1 + k\, (m_u/d_{ur})] \qquad (1a)$$

Written in this way, (1a) is the underlying merchandising model. Its form is preserved as we aggregate from individual store to market by means of ACV projection, as shown in (3a).

4. Product Line

Next we investigate the implications of (3) for a product line or other aggregate and see whether merchandising measures can be found that will model product line sales in a natural way. Let

$P \; = \; \{u\} \; =$ a set of UPC's that form a product line P.

$s \; = \; \Sigma_{u \in P}\, s_u \; =$ sales rate of product line P in the market (equivalent units/week).

$s_0 = \Sigma_{u\epsilon P}\ s_{0u}$ = base sales rate of product line P in the market (equivalent units/week).

$v_0 = \Sigma_{u\epsilon P}\ v_{0u}$ = base sales at full distribution of product line P in the market (equivalent units/week).

Within a product line different UPC's are likely to have quite different levels of distribution. Depending on the tastes and interests of particular retailers and the general popularity of individual items, certain sizes and packages will be found in some stores but not others. This suggests that, in constructing product line measures, a good weighting for individual components will be v_{0u}, base sales at full distribution. It takes account of the inherent strength of the product independent of its actual distribution.

Using v_{0u} as a weighting factor, we define aggregate measures for distribution and merchandising of the product line. Let

$$m = \Sigma_{u\epsilon P}\ (v_{0u}/v_0)\ m_u$$

= merchandising for product line P.

$$d = \Sigma_{u\epsilon P}\ (v_{0u}/v_0)\ d_u$$

= distribution for product line P.

Algebra reduces this last to:

$$d = s_0/v_0 \qquad (4)$$

Starting from (3) for the individual UPC's at market level, namely,

$$s_u = s_{0u}[1+k(m_u/d_u)] \qquad (3)$$

we sum over the u's in the product line:

$$s = \Sigma_u\ s_u$$

$$= \Sigma_u\ s_{0u} + k\ \Sigma_u\ s_{0u}\ (m_u/d_u)$$

$$= s_0 + k\ \Sigma_u\ v_{0u}\ m_u$$

$$= s_0 + k\ v_0\ m$$

$$s = s_0[1+k(m/d)] \qquad (5)$$

Therefore, <u>with the definitions we have used</u> for merchandising and distribution of the product aggregate, <u>the response model for the product</u>

line, (5), is the same as for the individual UPC, (3).

To summarize in words:

	Market level: product line		
	sales = (base sales) (1 + k [merchandising / distribution])		
where	merchandising	=	weighted merchandising over products
	distribution	=	weighted distribution over products
and	weight	=	(base sales) / (distribution) for each product.

As in the case of (3) for individual UPC's, we can express (5) in terms of the more fundamental base sales at full distribution, v_0:

$$s = v_0 \, d \, [1 + k \, (m/d)] \qquad (5a)$$

5. Aggregations of aggregations

Without developing a formal notation, we argue for the recursiveness of our aggregation process:

> Define an aggregation of product-lines, in which sales is the sum of individual product-line sales, and merchandising and distribution measures are constructed by weighting product-line measures by base product-line sales at full distribution. The resulting aggregated variables will satisfy equations analagous to (5) and (5a).

This follows from the steps that take (3) and (3a) into (5) and (5a), since the particular names of the atomic entities and aggregated entities are immaterial to the argument.

6. Time aggregates

The basic rhythm of a supermarket is weekly. Many customers are in the habit of shopping once a week. Feature advertising appears weekly. Special displays are put up and, although some may last longer than a week, weekly cycles pace the store's planning. For all these reasons the data companies collect and report store data with weeks as the finest time unit.

Although merchandising within a given week will often have effects beyond that week because customers may stock up on a product, this phenomenon is partially counteracted by normal purchase cycles, which for most products are several weeks or more. Although one set of customers may buy a product in a given week, a different set will usually buy it in the following one. In addition there is a tendency for retailers to separate multiple promotions of the same product with periods of other activity.

As a result, the first order effect of merchandising is to produce a separate response in each week that it takes place. This implies that an important measure of the aggregate effect over a multiple week period will be a simple sum of the activity in individual weeks. We can model this by adding a week subscript, t, to our basic equations, for example (3):

$$s_{ut} = s_{0ut} \, [1 + k \, (m_{ut}/d_{ut})] \qquad (6)$$

Typical time aggregations of interest would be four weeks or twelve weeks. In such periods we would not expect much variation in either distribution or base sales but we could easily encounter large week to week changes in merchandising and actual sales. Consequently, let $s_{0ut} = s_{0u}$ and $d_{ut} = d_u$, at least for the weeks within a given aggregate period. Let T index the aggregate periods, W be the number of weeks in one of them, and $t \epsilon T$ refer to the weeks in a specific T. Then, summing (6), the aggregate sales in T will be:

$$\Sigma_{t \epsilon T} \, s_{ut} = s_{0u} \, [W + k \, (\Sigma_{t \epsilon T} \, m_{ut})/d_u]. \qquad (7)$$

Here we see that incremental sales due to merchandising accumulate with ACV-weeks of merchandising (i.e., Σm_{ut}), the constant of proportionality being k times $s_{0u}/d_u = v_{0u}$, the base sales at full distribution.

Equation (7) can be put in a form analogous to earlier models by defining s_{uT} and s_{0uT} as aggregate sales and base sales respectively, and $m_{uT} = (\Sigma_{t \epsilon T} \, m_{uT})/W$, the average merchandising/week during the aggregate period. Then

$$s_{uT} = s_{0uT} \, [1 + k \, (m_{uT}/d_{uT})] \qquad (8)$$

which has the same form as (6).

7. Surrogates for base sales

Some companies using scanner data may not have the base sales measure, s_{0u}, at least not without extensive new analysis. We can still handle this case reasonably well. We do not wish to approximate s_{0u} by current sales, s_u, because this leads to using parts of current sales to predict current sales, a fallacious move. The important characteristics that a surrogate for s_{0u} should have are:

(1) It should be constant (or nearly so) over time so that it is just a scale factor as far as time series regressions are concerned.

(2) It should reasonably reflect the relative importance of each product in the volume of the product aggregate.

We shall not create spurious correlations if we satisfy (1) and we shall be giving reliable weights to each product if we satisfy (2).

We have tried various surrogates with quite good success. In the empirical testing to follow we created one as follows. First, an overall mechandising measure was constructed (see section on testing). Its values for all weeks under consideration were sorted by size. The weeks representing the lowest 50% of merchandising activity were then used to calculate an average sales rate under conditions of low merchandising. This became the surrogate s_{0u}.

8. Testing the Theory

The theory behind the new measures predicts the following outcomes:

(1) The response constant, k, calibrated for the product line using the new measures, will be similar to values of k calculated for the individual UPC's making up the line;

(2) The response constant for the line calculated from the old measures will be substantially different from that calculated from the new;

(3) Response constants calculated from the new multi-week time aggregations will be similar to those from single-week measures.

Our test bed is the set of Ocean Spray products in Table 1. The product line, Total Cranberry Juice Cocktail, is the sum of eight individual UPC's.

Many choices for a test merchandising variable are possible: displays, price-cuts, and newspaper features (which may be classified A ads, B ads, or C ads), either separately or in combination. After examining a number of regression runs of sales against these variables for specific UPC's, we have built a composite measure that is a weighted combination of all the variables. The weights are fairly arbitrary but do reflect the greater effectiveness of some merchandising activities relative to others. For present purposes, it is only important that same weights be used consistently in all the testing. The specific function is

merchandising = 0.64 (%ACV on display) + 0.32 (%ACV with A ads)
+ 0.16 (%ACV with B ads) + 0.08 (%ACV with C ads)
+ 0.04 (%ACV with price cuts)

Thus, if a UPC has 100% of ACV with displays, and A ads, and price-cuts, it will have merchandising = 100% (or 1.0).

We examine each prediction in turn.

8.1 Response constants are similar for UPC's and product line

The response constant, k, is estimated separately by regression for each UPC and again for the product line. The latter employs the new aggregate. We do two geographic cases, Total US, and New York. Table 2 presents the results.

As may be seen, although there are individual variations, the values are generally similar. This suggests that the aggregate measures we have constructed are doing their job and confirms the prediciton. A summary of the regressions behind Table 2 appears in the Appendix.

	Response constant, k	
Product	Total US	New York
Cran. Cocktail Liquid Conc.	10.1	7.4
Cran. Cocktail Aseptic 3-pack	6.4	3.6
Cran. Cocktail 32 oz. Bottle	4.9	5.1
Cran. Cocktail 48 oz. Bottle	4.6	5.3
Cran. Cocktail 64 oz. Bottle	6.2	4.7
Cran. Cocktail 128 oz. Bottle	9.5	4.7
Cran. Cocktail Low Cal. 32 oz.	10.1	5.0
Cran. Cocktail Low Cal. 48 oz.	4.0	4.5
Total Cranberry Juice Cocktail	7.6	4.3

Table 2. The merchandising response constant, k, differs somewhat by product within geographical area, as would be expected, but its value calculated from aggregate measures for the whole product line has a generally similar value. This is true for both New York and the US total.

8.2 New measures produce different answers from the old

We can run the same regressions to estimate the response constant, k, using the old measures of distribution and merchandising. The results for the individual items within the product line will be identical since the measures are exactly the same at the UPC level. Any change comes for the product line.

Table 3 shows that the old measures lead to quite different response constants. They fall outside the range of the constants for the individual UPC's by several standard deviations, thereby confirming the prediction. Essentially, the new aggregates are giving information about the product line that is analogous to that which the UPC-level regressions are reporting. The old aggregate measures are not. A summary of the regressions for Table 3 appears in the Appendix.

Total Ocean Spray Cran. Cocktail

Response constant, k
(standard errors in parentheses)

Merchandising Measures	Total US	New York
New	7.6 (0.6)	4.3 (0.3)
Old	2.0 (0.2)	1.0 (0.1)

Table 3. The new merchandising measures for the product line give response constants that are quite different from the old.

8.3 Multiple-week aggregates give similar results to single-week data

We have proposed that the multi-week aggregate of a merchandising measure be the simple sum of single week values. In the case of distribution an average is used.

Consider, for example, display. Table 4 shows that Ocean Spray 32 oz. Bottle has 3.2 % of ACV with display in the week of 2-22-87, 4.1 % the following week etc. For the 4-week period starting 2-22-87, the sum is 13.4 %ACV-weeks. We are proposing 13.4 as the aggregate display measure for the 4-week period.

Ocean Spray Cran. Cocktail 32 oz. Bottle

Week starting	Display (% ACV)	4-week total (%ACV-weeks)
2-22-87	3.2	
3-01-87	4.1	
3-08-87	3.6	
3-15-87	2.5	
		13.4
3-22-87	1.2	
3-29-87	1.3	
4-05-87	1.0	
4-12-87	1.0	
		4.5

Table 4. Merchandising measures for a multi-week period are simple sums of single week values.

The response model is now calibrated twice, once on weekly data (88 weeks) and again on 4-week aggregates (22 quadweeks). The response constants, k, are compared. Table 5 shows that the results are similar, as the theory would predict. We would not expect them to be identical since aggregation smooths the data and removes some of its natural variation.

	NEW YORK Market Response constant, k	
Product	weekly data	4-week data
Cran. Cocktail Liquid Conc.	7.4	8.1
Cran. Cocktail Aseptic 3-pack	3.6	3.1
Cran. Cocktail 32 oz. Bottle	5.1	6.2
Cran. Cocktail 48 oz. Bottle	5.3	4.9
Cran. Cocktail 64 oz. Bottle	4.7	4.7
Cran. Cocktail 128 oz. Bottle	4.7	2.1
Cran. Cocktail Low Cal. 32 oz.	5.0	6.0
Cran. Cocktail Low Cal. 48 oz.	4.5	4.3
Total Cranberry Juice Cocktail	4.3	3.8

Table 5. The merchandising response constants, k, are similar whether calculated from single-week or 4-week data.

8.4 Units of merchandising activity

The units of the time aggregates can be described as "% ACV-weeks." It is worth noting, however, that there is a strong analogy between these measures and "gross rating points" as used in television advertising. 100 rating points for a commercial in a time slot indicates that everyone in the geographic area under consideration has their sets tuned in to the program and an opportunity to see the commercial. Similarly, 100 % ACV display in a week means that all customers in the stores within the relevant geographical area have an opportunity to see a display.

Therefore, by analogy, we could say that the 32 oz. Bottle had 3.2 "display points" in the week of 2-22-87, and 13.4 "gross display points" for the four week period starting on the same date. Similar "gross feature points" and "gross price-reduction points" are easily defined. Such a terminolgy provides a useful alternative and synonym for "%ACV-weeks."

9. Conclusions

In order understand and discuss products and markets at a macro level, managers and market analysts need aggregate measures of sales, distribution and merchandising. Aggregation must be possible over time, geographic areas, and products. Good measures of merchandising and distribution are ones that are intuitively meaningful and drive sales.

We have motivated a new class of aggregate measures by hypothesizing a simple merchandising model at the UPC and store level and analytically aggregating it up to market and product line levels. The model makes sales propotional to distribution and linear with merchandising and is shown to retain its algebraic form at the aggregate levels if the measures are appropriately defined.

Since the UPC level response constant appears in the product line response model, we can estimate it on live data for both cases and compare the results. The values would not be expected to be identical but should be similar and they are. By way of contrast, estimates of the constant using currently used measures are different and inconsistent. In the same spirit, proposed new methods for time aggregation are found to produce similar values of the constant from weekly and 4-weekly data.

We do not suggest that our simple model captures all merchandising phenomena of marketing interest or that individual stores and UPC's all respond the same way. On the contrary, individual differences glossed over by the model offer many opportunities for improved marketing performance. What we do suggest, however is that the model captures the first order forces on sales produced by merchandising and distribution so that measures based on the model offer valid guidance at aggregate levels of product line and time periods.

The proposed aggregation rules summarize as follows:

Aggregate over:	Sales	Merchandising	Distribution
Time periods	simple sum	simple sum	simple average
Geographic areas	simple sum	weighted average weight = ACV	weighted average weight = ACV
Products	simple sum	weighted average weight = base sales at full distribution	weighted average weight = base sales at full distribution

Weighting by ACV is weighting by store or area size. Weighting by base sales at full distribution is weighting by the intrinsic strength of the product, i.e., how it sells when it is in the stores but has no special merchandising. This quantity has an intuitive appeal as a fundamental characteristic of a product or product line. It makes sense, when assessing, say, the quality of distribution for an overall brand, to weight more heavily the distribution of its strongest component sizes as measured in this way.

Finally, we find it quite striking that, when merchandising and distribution are defined as proposed, the merchandising model is recursive by level of product aggregation. In other words, it takes on the same analytic form at store and market levels, both for individual items and a product line, and the recursion can be continued into aggregations of aggregations.

REFERENCES

Abraham, Magid M. and Leonard M. Lodish (1987), "PROMOTER: An Automated Promotion Evaluation System," Marketing Science, 9 (Spring), 101-23.

IRI (1988), Infoscan Measure Selection Guide, Information Resources, Inc., 150 N. Clinton, Chicago, IL 60606 (December).

APPENDIX

A.1 Regressions using new measures

The database consists of 88 weeks (from the week starting 1-11-87 through that starting 9-11-88) for eight Ocean Spray Cranberry Juice Cocktail and their product line total. All products are expressed in equivalent units of 192 fluid ounces, with concentrated products as reconstituted equivalents. To remove seasonality the analysis was based on share of the juice category rather than sales. Base share for a product was taken to be the average share during time periods that had mechandising below its median value for the time series. The regression equation for each product is

$$(\text{share})_t/(\text{base share}) = a_0 + k\,(\text{merchandising})_t/(\text{distribution})_t + a_1\,(\text{week} - 44.5)_t/52$$

The purpose of the last term is to remove any time trend, since visual inspection shows that one sometimes is present. By centering it on the mean value of week, we remove its constant from a_0, which therefore has a model predicted value of 1.0.

------------------MARKET---------------------

-------------------TOTUS---------------------

PRODUCT	k	t-stat	a_1	t-stat	a_0	t-stat	R^2
OS CRN CKTL LQC	10.09	5.50	-0.12	-3.82	0.94	31.77	0.32
OS CRN CKTL 3PK	6.39	9.68	-0.03	-1.82	0.91	55.92	0.52
OS CRN CKTL 32	4.92	5.17	-0.07	-3.99	0.99	88.93	0.42
OS CRN CKTL 48	4.57	16.23	-0.11	-6.25	0.87	49.77	0.78
OS CRN CKTL 64	6.22	23.49	0.00	-0.08	0.91	50.78	0.87
OS CRN CKTL 128	9.46	8.58	-0.16	-10.57	0.99	102.24	0.73
OS C CKTL LC 32	10.09	9.85	-0.18	-12.43	1.02	141.30	0.84
OS C CKTL LC 48	3.95	9.14	0.00	0.05	0.97	88.55	0.52
OS CRN CKTL TOT	7.57	12.42	-0.07	-5.32	0.87	42.66	0.66

------------------MARKET---------------------

--------------------NY-----------------------

PRODUCT	k	t-stat	a_1	t-stat	a_0	t-stat	R^2
OS CRN CKTL LQC	7.41	9.14	0.07	1.78	1.02	42.96	0.52
OS CRN CKTL 3PK	3.62	11.52	0.12	4.56	0.97	61.58	0.65
OS CRN CKTL 32	5.11	6.16	0.06	2.86	0.99	94.94	0.30
OS CRN CKTL 48	5.29	18.65	0.10	3.03	0.91	42.18	0.80
OS CRN CKTL 64	4.69	28.25	0.17	6.11	1.01	63.37	0.91
OS CRN CKTL 128	4.72	6.02	-0.01	-0.47	0.98	85.83	0.29
OS C CKTL LC 32	5.00	6.87	-0.03	-1.53	1.01	92.76	0.41
OS C CKTL LC 48	4.53	11.23	0.25	7.41	0.93	48.60	0.63
OS CRN CKTL TOT	4.30	14.20	0.11	7.36	0.99	81.92	0.75

The only comment we might add to those in the text is that a_0 does indeed usually come out close to 1.0.

A.2 Regressions using old measures

The only run that is different is the product line:

MKT	PRODUCT	k	t-stat	a_1	t-stat	a_0	t-stat	R^2
TOTUS	OS CRN CKTL TOT	2.01	12.97	-0.07	-5.33	0.77	35.47	0.67
NY	OS CRN CKTL TOT	1.01	12.67	0.10	6.20	0.92	63.85	0.71

A.3 Regressions with different time periods

The 88 weeks become 22 quadweeks indexed by T. We use the model in the form

$$(\text{share})_T/(\text{base share}) = a_0 + k\ (\text{merchandising})_T/(\text{distribution})_T + a_1\ (\text{quadweek} - 11.5)_T/13$$

where share, merchandising and distribution are four week averages.

	MKT						
	NY						
PRODUCT	k	t-stat	a_1	t-stat	a_0	t-stat	R^2
LQCON	8.10	4.52	0.06	1.05	1.01	23.76	0.53
ASEP	3.06	4.55	0.12	3.74	0.99	38.23	0.64
R32OZ	6.18	3.44	0.07	2.16	0.98	62.19	0.33
R48OZ	4.93	12.95	0.10	3.06	0.93	37.68	0.89
R64OZ	4.66	16.13	0.17	4.52	1.01	43.33	0.94
R128OZ	2.10	1.50	-0.02	-0.78	1.00	69.78	0.06
LC32	6.03	4.59	-0.02	-0.83	1.00	75.37	0.59
LC48	4.34	4.33	0.25	4.11	0.94	24.38	0.55
CCKTL	3.80	7.48	0.11	6.84	1.01	56.08	0.84